RAW SEQUENCE LISTING
DATE: 09/04/2003
PATENT APPLICATION: US/09/804,409B
TIME: 09:36:14

Input Set : **N:\paola\US09804409B.raw.txt**
Output Set: **N:\CRF4\09032003\I804409B.raw**

```
 1 <110> APPLICANT: ENGENE, INC.
 2       KIEFFER, TIMOTHY J.
 3       CHEUNG, ANTHONY T.
 4 <120> TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATED PROTEIN
 5       EXPRESSION IN GUT
 6 <130> FILE REFERENCE: 029996/027 8721
C--> 7 <140> CURRENT APPLICATION NUMBER: US/09/804,409B
 8 <141> CURRENT FILING DATE: 2001-03-12
 9 <160> NUMBER OF SEQ ID NOS: 19
10 <170> SOFTWARE: PatentIn Ver. 2.1
12 <210> SEQ ID NO: 1
13 <211> LENGTH: 19
14 <212> TYPE: DNA
15 <213> ORGANISM: Artificial Sequence
16 <220> FEATURE:
17 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
18 <400> SEQUENCE: 1
19       ccagccgcag cctttgtga                                              19
21 <210> SEQ ID NO: 2
22 <211> LENGTH: 22
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
27 <400> SEQUENCE: 2
28       ggtacagcat tgttccacaa tg                                         22
30 <210> SEQ ID NO: 3
31 <211> LENGTH: 19
32 <212> TYPE: DNA
33 <213> ORGANISM: Artificial Sequence
34 <220> FEATURE:
35 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
36 <400> SEQUENCE: 3
37       accaccagcc ctaagtgat                                             19
39 <210> SEQ ID NO: 4
40 <211> LENGTH: 22
41 <212> TYPE: DNA
42 <213> ORGANISM: Artificial Sequence
43 <220> FEATURE:
44 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
45 <400> SEQUENCE: 4
46       ctagttgcag tagttctcca gc                                         22
48 <210> SEQ ID NO: 5
```

```
49 <211> LENGTH: 1319
50 <212> TYPE: DNA
51 <213> ORGANISM: Mus musculus
52 <400> SEQUENCE: 5
53      ccgaaattac ccactacgtt ggaattctat aagggttggg tttgctgttt tgtttacagc   60
54      tgcgtctttg gcacccagca cagctgagtg gttctaagcc cacgtcgatg cttaacacat  120
55      ggttgttgaa tgaatacacg cgaagccggt tctcatttag gggcatgagt aggcagaggt  180
56      gtgggcagga agcaggaaag agcggaaaca ggtgcggaca gaaaggaggg gctctgaagg  240
57      atgccagtca gtgccaaact gtcatccaga taccaggttc actgtggccc taggccaggc  300
58      tgcacggggc ttcccatgtg gtctgcccag ggtgagagca gaactgcggt gggcggggca  360
59      gaaggaaacc aaccaggaag cagggttgca cccaaattat ccaggtttta agtacattta  420
60      agagacaagg ctgggctgtt gaaggtcaga ggtgtccctg gggtgctgga ctaggactga  480
61      ccacttctgt tttagtttaa tggtgagaac tgcctcacac tgctacctgc cttacttgcc  540
62      ccttgagagc tgtgagccta ggacccaccc atgtgtgggt tggaccttca gtcacacact  600
63      gaacgtgtgt gaagccactg gttgtcagag cagggctctc ggcactgagg aagcagtgac  660
64      cactatcccc tatcaaataa caattaaata cacacagaat gcgaggcaca caactgagtt  720
65      tcaggagagg cctcgctcag gcaaggggtt caagaggctt ctgtgggacc cgctggatgt  780
66      tccagggagt tcttaaagat gggcgtgcct ccagccaagt gaaatcaaga gaaaagtacg  840
67      cgaagtatag gaaaactcag cagtctggag aggtaaatag gggaggaatc cgaggctcag  900
68      agacaggagt gacttgccca cggacgcaca gcaagttggc aggtggagtt cagctgtgcc  960
69      accttctgaa gccgggtacc ctttacagcc accagataca agcgggatag agacagctga 1020
70      tggagaagct ggaggtgggg ggcgggaccc cgaaggtggg gaaagggcgc ggggggggcgg 1080
71      tcctatgacg taatttcctg ggtgtgtgcg cgcgtgtgcg tgcgtgtgcg tgtatataaa 1140
72      agccggcata gcattgctgc tgctgccgcc gccaccgcca ccatcaccgc tgttaccacc 1200
73      accgctactg cagtgttccc gctggtgcag agctttggta gccagactac agacccactc 1260
74      ccgccatcct cctgcagcag ctcgtccact ctttccgcac cgtccggctc gctatgcgc  1319
76 <210> SEQ ID NO: 6
77 <211> LENGTH: 1760
78 <212> TYPE: DNA
79 <213> ORGANISM: Mus musculus
80 <400> SEQUENCE: 6
81      gggaactttc tctagctctt tcattagggg ccctgtgttc catctaatag ctgactgtga   60
82      gcatccactt ctgtgcttgc caggcactgg catagcctca caagagacag ctatatcagg  120
83      gtcttgtcag caaaatcttt ctggcatatg caatagtgtc tgggtttggt ggttgtatat  180
84      gggctggatc cccgggtggg gcagtctctg gatggtcttt ccttccgtct tagctccaaa  240
85      ctttgtctct gtaactcctt ccatgggtac tttgtttccc attctaagaa ggagcaaagt  300
86      atccacactt ccttcttctt ccttcttctt gagttttgca aatgccacaa aactttcaaa  360
87      gccttctgaa tagccttctc tttagtgctt tccaatgtat attaaaataa tctatctttc  420
88      atccccattg attaaagcct tcttaaagcc agaaaactat attcattttt ttcttttccc  480
89      agtagttcac aaactatctg gcacctcata agcatcataa ctcagttggt gggtagataa  540
90      aattggaatg tgattgttca gtcagcagag acttttagag gacctcatac aacaagattc  600
91      tctcagttct cagaaatata tttcagtata tacagggtta gaggactcac atctttaata  660
92      aaataaagtt aaaaatttag acctgtataa attattaagg tacctaatca agttccacgg  720
93      caaagtacag ccatggttat gaattataaa tccaagaagc ggtgggttaa ctctgacatt  780
94      gttccttgga tggttctcat tcattgaagt tagtcacctc aacttactca accaaaacct  840
95      agaagtattt ctgtggtact atgttctctt gatgccaaga gggctctagg catatgaaaa  900
96      tctctcaatc tctctccctc tctctccccc ttccacccccc actctctctc ttctagcagt  960
97      aatccctccc ttcctggtag gcagtatgtt ttttggagca cagtttctta gctatctctt 1020
98      gcaacacctg attttgctga agatttgaat ggcctcatat agaagtatca acaacttgag 1080
```

RAW SEQUENCE LISTING   DATE: 09/04/2003
PATENT APPLICATION: US/09/804,409B   TIME: 09:36:14

Input Set : **N:\paola\US09804409B.raw.txt**
Output Set: **N:\CRF4\09032003\I804409B.raw**

```
 99      cgtctgtgaa ctctcatttt gacactgtgc tgaaagaatt ggagttgatt ctcattaaaa 1140
100      aaaaaattaa gcatctcacc tttttgctc  aaactaaaca gttttaaaac agttctgcct 1200
101      ggagtcatga tatgaaatac gatctatcat atttgcaatg ttctgttcaa ttgtggctgc 1260
102      accaggaaat gagaagctat ttctttatag gcacaaataa aaagatagtc attatctgta 1320
103      aaattcttat gacatggcag caagcccaag aaacctttct aaacaaggcg tgaaaacgca 1380
104      gagatgtcct tgcaattagt catgtctatc tgacagattt cttcctttct aagggaattt 1440
105      gtgctgaaca ttttatttcg agcctcagag ataaaagaag ggggaagaag ctgtagtttt 1500
106      tgctacataa gacaggtggc gtaagcatgc aacgctttaa aaaaatatct aaagtgattg 1560
107      ttttctctcg gattctttga aaaagctcgc ctgcgctggg gtttgaggct gagccggtga 1620
108      cgtcagcgtg gaatgcggag tcaggcgccc aggctctcta taagccgagg agctgtccgg 1680
109      tgctgaaacg gcccgagccc tcactcagcg gcagagagga gcatgcttgg agccttccac 1740
110      ataatataag acagaggtaa                                             1760
```

112 <210> SEQ ID NO: 7
113 <211> LENGTH: 2623
114 <212> TYPE: DNA
115 <213> ORGANISM: Mus musculus
116 <400> SEQUENCE: 7

```
117      agctttaggt gtgtgaatat ctactttggt gctagggcct tggtcatact aagtaagttt 60
118      ccccttcact ggggtgtacc agtttaccct ggactgtcta agcaacaaga aggatagaca 120
119      tggcctacca cagatttcat gtctgccact ggctatgtca gaacatgtag gagcttttgg 180
120      aatcagtgaa acaggtattt tcagactgcc ttccctgcgt ggggctttcc cgaagccata 240
121      tttttcctag agtcagcctt tcccagctga ggacaagctg tactggacag atgccagcca 300
122      cttgaactgg gaatacatgg tcatttaggc agctggctta tctcatccat ggtacttgat 360
123      ggcttcgggt cagcacctca cagaaagttc agacgggagg cttccgagaa aacagagaag 420
124      caggcaggag atcctgcagg caatcctcct gctccacagc ctgcatggac ttccctcagc 480
125      cttagtgcgt gtgggtccca tctgagaaca ttggttatat gttattttca aaccgatctg 540
126      cctttaagga gtggaagaaa aaaactgtgg tgtttgggct acctttatga taatggcctt 600
127      ttcatcctcc taataaatat tgccaagtag ggtagattct atacgaaagc tcttaaccca 660
128      tggtattagc aaatcatgta ggtgctaata atgaatactg gatgcagtca gtacagggat 720
129      ataaaatgga atgtaagagc ctgttgctat gaatggttag ctaactagat gttgtacaag 780
130      aaatgttgac gttatgacgt gtggaaactt ggtattgaag atgtggactc gaaactttgt 840
131      ggattttttg atgccatgat aaaaatgtga agaatactgt tccttaccaa aaagaagaag 900
132      aagaaggaga aggaggagga agaggaggag gaggaagaag aggggagga  agaagaagag 960
133      aaggaggagg aagaggagga ggaggaagaa gaggaggagg aggaagaaga agagaaggag 1020
134      gaggactagg aggaggagga gaagaaggag aaggggaagg agagagtagc cagaacattt 1080
135      ggggtgccat cagaatacca gatactccag acatagtcac agaaggactg gtttgtttgt 1140
136      taaataggtg cttgaaaag  tttgtgggga aacctgcagt gagattgtgt gtcttagaaa 1200
137      tgataggcaa gattcatcca caagaatgcg acaagatggc tgcctgaaca agccctgaac 1260
138      attaacagca ccagtagacc tgcttacacg gaagaaagca atctcatagg ccctcacccc 1320
139      aaacaaagac tacagacagc agaggaactg gagagcagga gaaattgggt ctccctttta 1380
140      tgagccccct aactggttgt caaatactca atggtcagcc ctgaaatcat atgcacaaag 1440
141      taatactagc gcaactgaac agattgtagc tgtgtgtgt  tgtgtaatga taacaaagaa 1500
142      gaaaaggccc catgttagag agggagcaag gtgggcatgg aggtatggaa ggagttggaa 1560
143      ggaggggtga gaaggggaaa gtgatgtaat tatctttaa  tttataaaaa aataaaaaat 1620
144      gggctggtga gatggctcag tgggtaagag cacccgactg cttcttccga aggtctggag 1680
145      ttcaaatccc agcaaccaca tggtggctca caaccatccg taacgagatc tggcgccctc 1740
146      ttctggagtg tctgaagaca gctacagtgt acttacatat aataataaa  taaatctttt 1800
147      aaaaaaaata aaaataaaa  tattagaata aaatgtagag gaatatttt  aatttaacaa 1860
148      cttgggtgtg gcaaaagctt tcttcaacaa aaacttaatc cctcagataa gaaaagacta 1920
```

RAW SEQUENCE LISTING   DATE: 09/04/2003
PATENT APPLICATION: US/09/804,409B   TIME: 09:36:14

Input Set : **N:\paola\US09804409B.raw.txt**
Output Set: **N:\CRF4\09032003\I804409B.raw**

```
149        gaatccacga cgtggataga tacttctgta tgatgcaaga cactatttat caggttgtaa 1980
150        cttgagcaga acttgagttg taacttgttg ggaaacacaa cacccttggc aaacaaaaga 2040
151        ttactagata ttttagatga aatataaaaa tactttccac aactgatagg taggaaacag 2100
152        ttcaatagta atataattat tgaacaaata atccttaaaa gaagaaatcc agaggaatag 2160
153        caagttaggg gaagagaggg tgtgtgtgtg tgtgtgtgcg cgcacattta tagccaaaat 2220
154        agatgatata cttaaatgaa catgccatta aaacccatta ttttgcatac agtttacata 2280
155        tgctaatgaa tacttaaaaa aaaaacattg ggattggaga gaaatggctc agtggttaag 2340
156        agttcaattc ccagcaacca catgattgct cacaaccatc tgtaatggga tctgatgcct 2400
157        tcttctggta tgtctgaaga aagtgaccgt gtacttataa ttataaataa ataaatcttt 2460
158        aaccaaaaaa cccccataat ttcaacaaca gatatgtcct ggtctgaggc ttccaggcat 2520
159        agaaatagaa acacacagag tgtggagcca gtgcggttca ggtccgccat tccagttcag 2580
160        gcttcagacc aagagaaagg gaaaagaaga gacaagcaac aag                     2623
162 <210> SEQ ID NO: 8
163 <211> LENGTH: 226
164 <212> TYPE: DNA
165 <213> ORGANISM: Homo sapiens
166 <400> SEQUENCE: 8
167        tccaggaaat gcgcgatcca ggccggcggg cggggcgggg gctccggcga gaggggcgggc 60
168        cccgggaacg gcggcgggcg gggcgggagg cggggcccgg cccgttaaga agagcgtggc 120
169        cggccgcggc caccgctggc cccaggaaa gccgagcggc caccgagccg gcagagaccc 180
170        accgagcggc ggcggaggga gcgacgccgg ggcgcacgag ggcacc                   226
172 <210> SEQ ID NO: 9
173 <211> LENGTH: 110
174 <212> TYPE: PRT
175 <213> ORGANISM: Homo sapiens
176 <400> SEQUENCE: 9
177        Met Ala Leu Trp Met Arg Leu Leu Pro Leu Leu Ala Leu Leu Ala Leu
178          1               5                  10                  15
179        Trp Gly Pro Asp Pro Ala Ala Ala Phe Val Asn Gln His Leu Cys Gly
180                       20                  25                  30
181        Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe
182                  35                  40                  45
183        Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu Gln Val Gly
184             50                  55                  60
185        Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu
186        65                  70                  75                  80
187        Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys
188                       85                  90                  95
189        Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
190                  100                 105                 110
192 <210> SEQ ID NO: 10
193 <211> LENGTH: 450
194 <212> TYPE: DNA
195 <213> ORGANISM: Homo sapiens
196 <400> SEQUENCE: 10
197        gctgcatcag aagaggccat caagcacatc actgtccttc tgccatggcc ctgtggatgc 60
198        gcctcctgcc cctgctggcg ctgctggccc tctggggacc tgacccagcc gcagcctttg 120
199        tgaaccaaca cctgtgcggc tcacacctgg tggaagctct ctacctagtg tgcggggaac 180
200        gaggcttctt ctacacaccc aagacccgcc gggaggcaga ggacctgcag gtggggcagg 240
```

```
201        tggagctggg cgggggccct ggtgcaggca gcctgcagcc cttggccctg gaggggtccc  300
202        tgcagaagcg tggcattgtg gaacaatgct gtaccagcat ctgctccctc taccagctgg  360
203        agaactactg caactagacg cagcccgcag gcagcccccc acccgccgcc tcctgcaccg  420
204        agagagatgg aataaagccc ttgaaccagc                                    450
206  <210> SEQ ID NO: 11
207  <211> LENGTH: 167
208  <212> TYPE: PRT
209  <213> ORGANISM: Homo sapiens
210  <400> SEQUENCE: 11
211        Met His Trp Gly Thr Leu Cys Gly Phe Leu Trp Leu Trp Pro Tyr Leu
212         1               5                  10                  15
213        Phe Tyr Val Gln Ala Val Pro Ile Gln Lys Val Gln Asp Asp Thr Lys
214                        20                  25                  30
215        Thr Leu Ile Lys Thr Ile Val Thr Arg Ile Asn Asp Ile Ser His Thr
216                    35                  40                  45
217        Gln Ser Val Ser Ser Lys Gln Lys Val Thr Gly Leu Asp Phe Ile Pro
218                50                  55                  60
219        Gly Leu His Pro Ile Leu Thr Leu Ser Lys Met Asp Gln Thr Leu Ala
220         65                  70                  75                  80
221        Val Tyr Gln Gln Ile Leu Thr Ser Met Pro Ser Arg Asn Val Ile Gln
222                        85                  90                  95
223        Ile Ser Asn Asp Leu Glu Asn Leu Arg Asp Leu Leu His Val Leu Ala
224                    100                 105                 110
225        Phe Ser Lys Ser Cys His Leu Pro Trp Ala Ser Gly Leu Glu Thr Leu
226                115                 120                 125
227        Asp Ser Leu Gly Gly Val Leu Glu Ala Ser Gly Tyr Ser Thr Glu Val
228            130                 135                 140
229        Val Ala Leu Ser Arg Leu Gln Gly Ser Leu Gln Asp Met Leu Trp Gln
230        145                 150                 155                 160
231        Leu Asp Leu Ser Pro Gly Cys
232                        165
234  <210> SEQ ID NO: 12
235  <211> LENGTH: 3408
236  <212> TYPE: DNA
237  <213> ORGANISM: Homo sapiens
238  <400> SEQUENCE: 12
239        tctgttttca ggcccaagaa gcccatcctg ggaaggaaaa tgcattgggg aaccctgtgc  60
240        ggattcttgt ggctttggcc ctatcttttc tatgtccaag ctgtgcccat ccaaaaagtc  120
241        caagatgaca ccaaaaccct catcaagaca attgtcacca ggatcaatga catttcacac  180
242        acgcagtcag tctcctccaa acagaaagtc accggtttgg acttcattcc tgggctccac  240
243        cccatcctga ccttatccaa gatggaccag acactggcag tctaccaaca gatcctcacc  300
244        agtatgcctt ccagaaacgt gatccaaata tccaacgacc tggagaacct ccgggatctt  360
245        cttcacgtgc tggccttctc taagagctgc cacttgccct gggccagtgg cctggagacc  420
246        ttggacagcc tggggggtgt cctggaagct tcaggctact ccacagaggt ggtggccctg  480
247        agcaggctgc aggggtctct gcaggacatg ctgtggcagc tggacctcag ccctgggtgc  540
248        tgaggccttg aaggtcactc ttcctgcaag gactacgtta agggaaggaa ctctggcttc  600
249        caggtatctc caggattgaa gagcattgca tggacacccc ttatccagga ctctgtcaat  660
250        ttccctgact cctctaagcc actcttccaa aggcataaga ccctaagcct cctttttgctt 720
251        gaaaccaaag atatatacac aggatcctat tctcaccagg aaggggg tcc acccagcaaa  780
```

VERIFICATION SUMMARY   DATE: 09/04/2003
PATENT APPLICATION: **US/09/804,409B**   TIME: 09:36:15

Input Set : **N:\paola\US09804409B.raw.txt**
Output Set: **N:\CRF4\09032003\I804409B.raw**

L:7 M:270 C: Current Application Number differs, Wrong Format